



Identification d'erreurs de calcul de moyenne chez des étudiants de Master 2 en Sciences humaines et sociales.

Diane Diaz, Jean-Claude Régnier

► To cite this version:

Diane Diaz, Jean-Claude Régnier. Identification d'erreurs de calcul de moyenne chez des étudiants de Master 2 en Sciences humaines et sociales.. 4ème Colloque Francophone International sur l'Enseignement de la statistique, Société Française de Statistique - groupe enseignement de la statistique, Jan 2015, Bordeaux, France. pp.221-226. hal-01117064

HAL Id: hal-01117064

<https://inria.hal.science/hal-01117064>

Submitted on 16 Feb 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IDENTIFICATION D'ERREURS DE CALCUL DE MOYENNE CHEZ DES ÉTUDIANTS DE MASTER 2 EN SCIENCES HUMAINES ET SOCIALES.

Diane Diaz ¹ & Jean-Claude Régnier ²

¹ Université Lumière-Lyon2 ISPEF Sciences de l'éducation
86, rue Pasteur 69007 Lyon
diane.diaz@uni-lyon2.fr

² Université Lumière-Lyon2 ISPEF Sciences de l'éducation
86, rue Pasteur 69007 Lyon
jean-claude.regnier@univ-lyon2.fr

Résumé. Une notion aussi élémentaire que la moyenne arithmétique peut produire des erreurs et engendrer des difficultés auprès d'étudiants de sciences humaines et sociales. Il est admis que ce public dans sa majorité s'est forgé une opinion négative à l'encontre de la statistique et peut éprouver jusqu'à de l'aversion envers cette discipline (Gordon, 2004 p.46-47). Quelles sont alors les erreurs récurrentes et les principales difficultés rencontrées par les étudiants ? Pour répondre à cette question nous partons d'une enquête plus ample par questionnaire mise au point par Jean-Claude Régnier. De ces travaux nous nous limitons aux réponses récoltées de 2009 à 2011 auprès d'étudiants ayant atteint le dernier niveau d'études de second cycle qu'est le Master 2. Nous en extrapolons une situation qui décrit un événement quotidien et évalue des connaissances procédurales (Anderson, 1976, 1990 ; Tardif, 1992) à travers le calcul d'une moyenne arithmétique pondérée. Cette étude a donc pour objectif de caractériser les erreurs commises par les étudiants. L'analyse des réponses obtenues permet d'identifier des erreurs très variées. Les étudiants attestent qu'ils connaissent la méthode de calcul de la moyenne mais lorsqu'il s'agit de l'appliquer moins de la moitié d'entre eux y parvient. Les erreurs relevées sont de divers ordres, elles vont de l'incompréhension de l'énoncé à des erreurs de surface comme le report erroné des données. Il apparaît que la compétence de savoir lire un énoncé comme celle d'en extraire l'information pertinente ne soient pas suffisamment maîtrisées, même chez des sujets ayant atteint la 5^{ème} année d'études universitaires.

Mots-clés. Didactique de la statistique, pédagogie universitaire, apprentissage, difficultés.

Abstract. Something as basic as the arithmetic mean can produce errors and cause difficulties with students in the humanities and social sciences. It is recognized this public feels aversion with statistics and has gained a negative opinion against this discipline (Gordon, 2004 p.46-47). So what are the recurring errors and the main difficulties encountered by students ? To answer this question we start from a broader survey questionnaire developed by Jean-Claude Régnier. In this paper we limit ourselves to responses collected from 2009 to 2011 to students who are at the end of the second cycle of the Masters Course. We extrapolate a situation that describes an event of daily life and assess procedural knowledge (Anderson, 1976, 1990; Tardif, 1992) by calculating a weighted arithmetic mean. This study therefore aims to characterize the mistakes made by students. The analysis of the responses identifies a wide variety of errors. Students acknowledge awareness of the method of averaging, but when it comes to applying less than half of them succeeds. Errors found are of various kinds, they range from misunderstanding of the statement to surface errors like wrong reporting of data. It appears that the competence to read a statement or to extract the relevant information may not be sufficiently controlled, even in subjects who have reached university 5th year.

Keywords: Didactics of the statistics, university pedagogy, learning, difficulties.

1 Introduction

Même les concepts d'apparence les plus simples peuvent donner du fil à retordre et produire divers types d'erreurs. Aussi, quand il s'agit d'aborder une des premières notions de statistique qu'est la moyenne arithmétique il semble judicieux de s'interroger sur les difficultés et les obstacles que les étudiants devront surmonter pour la maîtriser. Combien d'étudiants peuvent résoudre un problème faisant appel à un calcul de moyenne arithmétique ? Quelles sont les erreurs récurrentes et les principales difficultés rencontrées par des étudiants de sciences humaines et sociales ? Nous nous intéresserons ici exclusivement à ce public dont il est aujourd'hui largement admis qu'il éprouve des difficultés, voire parfois de l'aversion, envers la statistique et s'est forgé une opinion négative envers cette discipline (Gordon, 2004, p.46-47) (Bruyninckx et al., 2001). Pour sa part Régnier (2003, p.179) relève la méconnaissance des notions de base en statistique des étudiants inscrits en sciences de l'éducation.

Les résultats présentés dans ce texte sont extraits d'une enquête par questionnaire plus ample mise au point par Jean-Claude Régnier en didactique de la statistique. A partir des réponses obtenues nous avons tenté de relever les erreurs les plus courantes commises dans la résolution d'une situation suscitant un calcul de moyenne arithmétique. Cela nous permettra d'établir des classes de difficultés rencontrées.

Si la moyenne arithmétique demeure l'un des concepts de statistique le plus basique il paraît intéressant (Batanero et al., 1994, p. 4) de l'étudier à plusieurs titres. Dans la vie courante, elle offre de nombreuses applications (ibid.) et permet de comparer facilement de grandes quantités de données. De plus, dans les revues scientifiques, la majorité des traitements se fondent sur des moyennes pour caractériser les données. La statistique inférentielle, de même, repose essentiellement sur des moyennes et leurs comparaisons (Pollatsek et al., 1981, p. 191). Or, malgré la familiarité avec ce paramètre d'apparence facile, un des paradoxes réside dans la difficulté que l'on rencontre dès qu'il s'agit de tenter de le définir.

Le concept de moyenne est plus complexe que ce qu'il n'y paraît. Il a d'ailleurs été validé que très peu d'étudiants parviennent à expliquer le choix de cet outil de calcul et à donner du sens au résultat qu'ils trouvent. Qu'ils soient en deuxième année d'IUT (Girard, 1998, p. 61) ou en formation des maîtres (Gattuso et Mary, 1997, 1999) les résultats de ces études font apparaître que la plupart des répondants se trouvent en difficulté.

2 Méthode

Depuis 2008, lors de la première séance de cours au mois de septembre de chaque année, un questionnaire est soumis à l'ensemble des étudiants en Licence (L3) et en Master (M1, M2) de Sciences de l'Éducation dans le cadre du cours d'initiation au raisonnement statistique et celui de statistique et traitement de données. Telle une évaluation diagnostique, ce questionnaire permet de récupérer des traces relatives aux connaissances et compétences et apporte des éléments pour identifier des facilités comme des difficultés rencontrées par les étudiants avant le début de leur formation.

Notre corpus se compose des réponses récoltées durant trois années universitaires allant de 2009 à 2011 et se limite au niveau d'études le plus élevé qu'est le Master 2. La taille de l'échantillon est restreinte, cependant elle couvre la presque totalité des étudiants présents à la formation observée. En effet, dans ce contexte de recueil en situation ordinaire, le mode de distribution de la main à la main, effectué au début d'un cours, permet d'obtenir un taux de retour exceptionnel de cent pour cent puisque tous les étudiants présents lors la première séance de cours remettent leur questionnaire complété. Cette étude longitudinale nous permet de regrouper trois microéchantillons comme le recommandent Marien et Beaud (2003, p.14) dans le cas de petits échantillons.

Ce sous-échantillon de la population a été choisi car il représente des étudiants qui ont dépassé le premier cycle, ont surmonté la sélection d'entrée en Master 2 et se trouvent en cinquième année

après le baccalauréat. Ce grade correspond au dernier diplôme universitaire encadré par des cours magistraux, des travaux dirigés ou des travaux pratiques selon des parcours collectifs prédéfinis selon un canevas modulable. En cela, nous nous démarquons de la plupart des enquêtes en didactique de la statistique universitaire qui s'intéressent essentiellement aux étudiants de premier cycle.

La situation proposée décrit un événement de la vie quotidienne et s'attèle à vérifier principalement des connaissances procédurales (Anderson, 1976, 1990 ; Tardif, 1992) à travers le calcul d'une moyenne arithmétique pondérée.

Situation 1

Un ascenseur dont la capacité maximale est de 700 kg, doit transporter dix personnes : quatre femmes dont le poids moyen est de 60 kg et des hommes de poids moyens quatre vingt kg. Ces personnes pourront-elles être transportées en un seul voyage ? Quel est le poids moyen de ce groupe de personne ?

La formulation repose sur une expression issue des mathématiques et non sur une énonciation provenant du paradigme classique de la statistique. Les études sur le calcul de la moyenne ne sont pas nouvelles. Nous avons pu trouver dans la littérature des résolutions de problèmes similaires (Pollatsek et al. 1981, p. 195 ; Mary et Gattuso 2005 ; Bihan-Poudec, 2010). Contrairement à ces exemples, la situation produite par Jean-Claude Régnier ne dénombre pas précisément les individus ce qui revient à poser une équation du premier degré à une inconnue. À titre indicatif, une des réponses pourrait être :

$$4 + n = 10 \text{ alors } n = 6 \text{ il y a donc 6 hommes}$$

$$(4 \times 60) + (6 \times 80) = 720$$

1. Non, ces personnes ne pourront pas être transportées en un seul voyage car le poids total est supérieur à la charge maximale ($720 > 700$).

2. Le poids moyen de ce groupe de personnes est de : $720/10 = 72$ soit 72 kg.

En résumé, cette situation requiert la compréhension de l'énoncé, le choix des calculs opérationnels, la capacité à effectuer les calculs et enfin l'aptitude à produire une réponse écrite.

3 Les erreurs identifiées

Sur les soixante-douze questionnaires récoltés, seuls trente-quatre -soit moins de la moitié des réponses- révèlent des commentaires conformes aux attendus. Le tableau ci-dessous permet de distinguer deux familles de réponses non conformes aux attendus :

- l'encodage « réponse fournie non opérationnelle » regroupe :
les erreurs de raisonnement, les erreurs de calculs, les erreurs de résultats, et les erreurs de prélèvement des nombres.
- l'encodage « réponse fournie partiellement » regroupe :
les réponses ne rapportant pas d'erreurs manifestes mais ne fournissant pas la totalité des résultats attendus.

	Absence de réponse	Réponse fournie non opérationnelle	Réponse fournie partiellement	Réponse fournie attendue
Situation 1	1	24	13	34

TABLEAU 1 *Réponses obtenues*

Nous constatons que vingt-quatre étudiants ne relèvent pas la totalité des données de l'énoncé, ils ne remarquent pas qu'ils disposent de suffisamment d'informations pour exécuter les calculs. Cinq répondants écrivent qu'ils ne connaissent pas le nombre d'hommes. Ils produisent alors une erreur issue de l'incompréhension de l'énoncé et ne semblent pas maîtriser l'extraction de l'information pertinente. Allant dans le même sens, onze étudiants produisent des calculs qui laissent présager la même déduction. En effet, sept d'entre eux procèdent à un ensemble d'opération arithmétique leur permettant d'obtenir le nombre d'hommes qu'il serait possible de transporter dans l'ascenseur dont la capacité maximale est de 700 kg, ce qui diffère nettement de la requête initiale. Onze étudiants fournissent la valeur 70 kg pour le poids moyen du groupe. Si huit étudiants ne détaillent pas le calcul qui leur a permis d'obtenir ce résultat, ils ont certainement calculé le poids moyen d'un homme et d'une femme, soit : $(60\text{kg}+80\text{kg})/2 = 70\text{kg}$, comme cela a été spécifié sur trois questionnaires. Dans ce cas, encore une fois, le nombre exact d'individus à transporter n'est pas pris en compte et les étudiants ne procèdent pas à un calcul de moyenne pondérée mais à une simple moyenne arithmétique. La stratégie de résolution est inappropriée. Cette absence du principe de la pondération se retrouve également chez des étudiants interrogés par Pollatsek et al. (1981).

De plus, trois étudiants n'ont pas prélevé les nombres exacts indiqués dans le questionnaire. Par exemple, certains n'ont pas reporté le poids des hommes spécifié dans l'énoncé. Un étudiant rapporte alors 90 kg au lieu de la valeur 80 kg inscrite. Cette erreur, si elle paraît faible et n'être sans doute qu'une erreur de surface, fausse cependant les résultats obtenus. Un étudiant va jusqu'à restituer 20 kg. Même si le trait d'union de quatre-vingt n'apparaît pas dans l'énoncé, l'étudiant aurait alors pu s'interroger sur la vraisemblance d'une si petite masse qui ne peut pas correspondre à celle d'un homme. Ces exemples révèlent une transposition incorrecte des données. Dans la même lignée un répondant change l'unité de mesure et indique que le poids moyen des individus est de 70 grammes ! Il arrive aussi pour douze étudiants de substituer le poids total au poids moyen et inscrivent que le poids moyen des individus s'élève à plusieurs centaines de kilogrammes (640, 700, ou 720). Cette confusion entre poids total et poids moyen confirme les observations cliniques produites sur un exercice similaire par Bihan-Poudec (2010, p. 90-92). Sa recherche donne lieu à des réponses écrites laissant percevoir l'amalgame entre les deux notions. C'est à la lecture interrogative à voix haute du résultat par le chercheur que les trois enfants consultés rectifient spontanément verbalement leurs imprécisions initiales.

4 Discussion

Alors même que l'énoncé est très descriptif et dépeint un événement concret de la vie courante, il est à noter qu'au total treize répondants ne tiennent pas compte de l'ordre de grandeur requis. Il semble qu'ils n'ont pas vérifié la conformité de leurs résultats, car s'ils l'avaient fait, ils auraient sans doute été alertés par la démesure de leurs conclusions. A cet égard, Girard (1998, p.61) signale combien le sens d'un résultat relève du contexte situationnel. L'erreur, due à une carence de recontextualisation, est également constatée dans les études menées par Jean-Claude Régnier (2003, p.13) et Gattuso (2005, p.85). Allant dans ce sens, Bihan-Poudec va même jusqu'à s'interroger sur le paradoxe suivant :

Tout se passe comme si plus l'apprenant maîtrisait les modalités de calculs, moins ces derniers étaient connectés à la réalité d'où les données sont issues. Nous pourrions bien déboucher sur un paradoxe : si la schématisation mathématique aboutit à une indépendance au contexte, elle débouche sur une situation où le traitement statistique se suffit à lui-même, se trouve déconnecté du contexte initial et surtout apparaît sans retour sur celui-ci, retour qui permettrait pourtant d'interroger la validité des résultats obtenus. Bihan-Poudec (2010, p. 92)

Rien ne nous permet de confirmer ou d'infirmer cette contradiction, cependant, nous pouvons remarquer que le contexte, issu de la situation décrivant un événement quotidien, n'est pas pris en compte par plus d'un cinquième des étudiants.

De plus, le calcul de moyenne arithmétique simple, sans l'ajout de pondération, nous évoque un pan de la théorie des champs conceptuels (Vergnaud, 1991, 1994, 1996) dont l'analyse de problèmes sur

l'addition met en évidence la notion de schème qui correspond à l'« *organisation invariante de l'activité pour une classe de situations données* ». Il semblerait que onze étudiants observés appliquent un schème généralement efficient dans le cas d'un calcul de moyenne mais sans tenir compte des spécificités propres au contexte particulier de l'étude qui exige l'ajout d'une pondération. En réponse à une question reposant sur un calcul de moyenne, ces étudiants procèdent l'opération qu'ils ont mémorisée et sont capables de reproduire une structure organisationnelle, c'est-à-dire un schème, mais sans l'adapter au contexte propre à la situation qui impose l'application d'un autre algorithme mathématique.

5 Conclusion

L'analyse des réponses obtenues a permis d'identifier des erreurs très variées. La plupart des étudiants connaît l'algorithme de calcul de la moyenne mais lorsqu'il s'agit de l'appliquer moins de la moitié d'entre eux y parvient. Ainsi nous avons pu repérer les erreurs suivantes :

- mauvaise compréhension de l'énoncé ;
- stratégie de résolution inappropriée ;
- absence de la prise en compte de la pondération dans la résolution du problème ;
- transposition incorrecte des données sans prise en compte du contexte ;
- confusion entre le poids total et le poids moyen ;

Les erreurs relevées sont de divers ordres qui vont de l'incompréhension de l'énoncé à des erreurs de surface comme le report erroné des données, qui peut alors être attribué à l'*esprit fatigué* pour reprendre l'expression de Bachelard. Cependant, si l'étudiant avait procédé à une rapide estimation des résultats obtenus, nous aurions pu espérer qu'il procède alors une autocorrection. Dans ces exemples une large part de la consigne n'a pas été respectée et il semble que la compétence de savoir lire un énoncé comme celle d'en extraire l'information pertinente ne soient pas suffisamment maîtrisées, même chez des sujets ayant atteint la 5^{ème} année d'études universitaires alors même qu'elles constituent une part importante de la réussite universitaire.

Pour donner suite à cette étude, dans nos prochains travaux, nous allons nous attacher à comparer les réponses obtenues sur la moyenne avec des questions abordant d'autres concepts de statistique. Ainsi nous espérons faire émerger les difficultés engendrées par l'apprentissage de la statistique pour proposer des remédiations adaptées.

Bibliographie

- [1] Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, 3(1), 40-59.
- [2] Bruyninckx, M., Couvreur, N., Landercy, A. (2001). Perceptions, attitudes, craintes et attentes des étudiants de première candidature en psychologie et en sciences de l'éducation quant à l'enseignement des statistiques. *Actes de 18^{ème} Congrès de l'Association Internationale de Pédagogie Universitaire*, Dakar, Sénégal, p. 17
- [3] Régnier, J. C. (2003). A propos de la formation en statistique. Approches praxéologiques et épistémologiques de questions du champ de la didactique de la statistique. *Revue du Centre de recherche en Education*, 22, 157-201.
- [4] Zendreda, N. (2004). Difficultés et obstacles rencontrés dans l'apprentissage des tests paramétriques: objets d'une recherche en didactique de la statistique in *Actes des XXXVI^{èmes} Journées de Statistique*. Société Française de Statistique.
- [5] Calmant, P., Ducarme, M., & Schneider, M. (2011). Obstacles a priori à l'apprentissage de l'analyse statistique inférentielle. *Statistique et Enseignement*, 2(1), 43-59.
- [6] Nimier, J. (1976). *Mathématique et affectivité: une explication des échecs et des réussites*. Paris : Stock.
- [7] Nimier, J. (1988). *Les modes de relations aux mathématiques: attitudes et représentations*. Paris : Méridiens Klincksieck.

- [8] Batanero, C., Godino, J. D., Vallecillos, A., Green, D. R., Holmes, P. (1994). Errors and difficulties in understanding elementary statistical concepts. *International Journal of Mathematical Education in Science and Technology*, 25(4), 527-547.
- [9] Pollatsek, A., Lima, S., Well, A. D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics*, 12(2), 191-204.
- [10] Girard, J. C. (1998). A bas la moyenne! ou A propos des paramètres de tendance centrale et de dispersion d'une série statistique. *Repères-IREM n 33, Ed. Topiques*.
- [11] Gattuso, L., Mary, C. (1997). La moyenne, un concept évident? *Bulletin AMQ*, 37(3), 10-19.
- [12] Gattuso, L. (1999). La moyenne: un concept inexploité, d'une richesse exceptionnelle. *Repères IREM*, 34.
- [13] Marien, B., Beaud, J. P. (2003). *Guide pratique pour l'utilisation de la statistique en recherche. Le cas des petits échantillons*. http://eprints.aidenligne-francais-universite.auf.org/379/1/Guide_pratique_pour_l_utilisation_de_la_statistique_en_recherche.pdf (consulté 29/11/2014)
- [14] Anderson, J. R. (1976). *Language, Memory and Thought*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- [15] Tardif, J. (1992). *Pour un enseignement stratégique: l'apport de la psychologie cognitive*. Montréal: Les Éditions Logiques.
- [16] Mary, C., Gattuso, L. (2005). Trois problèmes semblables de moyenne pas si semblables que ça ! L'influence de la structure d'un problème sur les réponses des élèves. *Statistics Education Research Journal*, 4(2). 82-102.
- [17] Bihan-Poudec, A. (2010). L'enseignement de la statistique: en premier lieu, l'apprenant. *Statistics Education Research Journal*, 9(2), 88-103.
- [18] Vergnaud, G., (1991) Théorie des champs conceptuels, *Recherches en Didactique des Mathématiques*, vol.10/2.3 p.133-169
- [19] Vergnaud, G., (1994) Le rôle de l'enseignant à la lumière des concepts de schème et de champ conceptuel., in M. Artigue, R. Gras, C. Laborde, P. Tavnignot (Eds) *Vingt ans de didactique des mathématiques* Grenoble : La pensée Sauvage, p.177-191
- [20] Vergnaud, G., (1996) Au fond de l'action, la conceptualisation, in J.M. Barbier (coord) *Savoirs théoriques et savoirs d'action*, Paris : PUF.